REMARKS

Claims 1, 34, 37, 40 and 43 have been amended. The amendments are fully supported in the application and original claims. Claims 2-33, 35-36, 38-39 and 41-42 were previously canceled. No new matter has been added. Subsequent to this amendment, claims 1, 34, 37, 40 and 43-47 are pending and at issue.

Also, as requested by the Office Action, the following sets forth related co-pending applications with claims directed to GDF-8 (myostatin) polypeptides or fragments thereof:

U.S. Application Serial No. 09/708,693 filed November 7, 2000;

U.S. Application Serial No. 10/335,483 filed December 31, 2002;

U.S. Application Serial No. 10/997,809 filed November 23, 2004; and

U.S. Application Serial No. 10/991,343 filed November 15, 2004.

I. Amendment to the Specification and Claims

Per the request of the Office Action, the co-pending applications referenced in the application have been updated. Both U.S. Serial No. 09/124,180, filed July 28, 1998 and U.S. Serial No. 09/019,070, filed February 5, 1998 are now abandoned. No new matter has been added.

Claims 1, 34, 37, 40 and 43 have been amended. Claims 1, 34 and 40 have been amended and the claimed "activity" deleted.

Claims 47 and 43 have been amended to again delete the phrase, "as set forth in". This phrase was initially deleted in the amendment filed in response to the Notice of Non-compliant Amendment mailed March 16, 2005. However, it is again deleted herein to further advance prosecution of the above application.

II. Double patenting rejection

Claims 1, 34, 37, 40 and 43 are provisionally rejected on the ground of nonstatutory obviousness-type double patenting as being unpatentable over claim 20 of copending Application No. 10/997,809, which is allegedly directed to GDF-8 from *chicken* corresponding to SEQ ID NO:8 of the present invention (page 3 of the Office Action).

It is submitted that claim 20 of U.S. Application No. 10/997,809, filed November 23, 2004, directed to a "substantially pure growth differentiation factor-8 (GDF-8) having the amino acid sequence as set forth in SEQ ID NO:21". SEQ ID NO:21 is a *bovine* sequence. The bovine sequence of the present invention is SEQ ID NO:12 (claims 1, 34, 37, 40 and 43). Hence, Applicants have submitted herewith a Terminal Disclaimer, disclaiming the term of any patent that issues from the subject application that may extend beyond the term of co-pending U.S. Pat. No. 10/997,809, filed November 23, 2004. Accordingly, it is respectfully requested that this rejection be removed.

III. Rejections under 35 U.S.C. §102

Claims 1, 34, 37, 40, 43, 46 and 47 are rejected under 35 U.S.C. §102(e) as allegedly anticipated by Barker et al. (U.S. Patent No. 6,369,201). Applicants respectfully traverse the rejection as it applies to the pending claims.

According to the Office Action, claims 1, 34 and 40 are directed to various activities which are not disclosed in the parent U.S. Application No. 09/124,180, filed July 28, 1998. Thus, according to the Office Action, Barker et al. filed on February 18, 1999 is prior art against the instant claims.

Applicants submit, that according to MPEP 2163.07(a) the skilled artisan would accept that the claimed promyostatin polypeptides had the inherent function at the time of the filing of the present application:

> By disclosing in a patent application a device that inherently performs a function or has a property, operates according to a theory or has an advantage, *a patent application necessarily discloses that function, theory or advantage, even though it says nothing explicit concerning it*...In re Reynolds, 443 F.2d 384, 170 USPQ 94 (CCPA 1971); *In re Smythe*, 480 F. 2d 1376, 178 USPQ 279 (CCPA 1973). "To establish inherency, the extrinsic evidence 'must make clear that the missing descriptive matter is necessarily present in the thing described in the reference, and that it would be so recognized by persons of ordinary skill. Inherency, however, may not be established by probabilities or possibilities. The mere fact that a certain thing may result from a given set of circumstances is not sufficient.'" *In re Robertson*, 169 F.3d 743, 745, 49 USPQ2d 1949, 1950-51 (Fed. Cir. 1999) (citations omitted).

It is submitted that at the time of the filing of U.S. Application No. 09/124,180, filed July 28, 1998, signal peptide domains were known to have signal peptide activity (claim 1). See Exhibit A, Nothwehr et al. (1989) Eukaryotic signal peptide structure/function relationships *J Biol Chem* 264(7):3979-3987. It is also submitted that both the promyostatin prodomain and myostatin domain (claims 34 and 40) were known to have myostatin binding activity, which activity inhibited muscle growth activity. See Exhibit B, McPherron & Lee 1997 Double muscling in cattle due to mutations in the myostatin gene *PNAS USA* 94:12457-61.
Thus, the skilled artisan would recognize that the activities of the claimed polypeptides were either well known in the art or inherent in the structure at the time of the filing of U.S. Application No. 09/124,180. Therefore, Applicants have a valid claim of priority to U.S. Application No. 09/124,180, filed July 28, 1998, which is before U.S. Patent No. 6,369,201 to Barker, filed on February 18, 1999. Accordingly, Barker cannot anticipate the claimed invention.

However, without conceding to the correctness of the rejection as discussed above, Applicants have amended claims 1, 34 and 40, and dependent claims therein, such that the claimed activities have been deleted, i.e. deletion of "having signal peptide activity"; "having myostatin binding activity"; and "having muscle growth inhibitory activity", respectively. Thus, the rejection is also moot with regards to claims 1, 34, 37, 40, 43, 46 and 47.

Accordingly, withdrawal of rejection of claims 1, 34, 37, 40, 43, 46 and 47 under 35 U.S.C. §102 is respectfully requested.

Conclusion

In view of the amendments and above remarks, it is submitted that the claims are in condition for allowance, and a notice to that effect is respectfully requested. The Examiner is invited to contact Applicant's undersigned representative if there are any questions relating to this application.

A check in the total amount of $125.00 is enclosed as payment for the One-Month Extension of Time fee ($60.00) and Terminal Disclaimer fee ($65.00). No other fee is deemed necessary with the filing of this paper. However if any fees are due, the Commissioner is hereby authorized to charge any fees, or make any credits, to Deposit Account No. 07-1896 referencing the above-identified attorney docket number. A copy of the Transmittal Sheet is enclosed.

Respectfully submitted,

Date: October 19, 2006

Lisa A. Haile, J.D., Ph.D.
Registration No. 38,347
Telephone: (858) 677-1456
Facsimile: (858) 677-1465

DLA PIPER US LLP
4365 Executive Drive, Suite 1100
San Diego, California 92121-2133
USPTO Customer No. 28213

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 264, No. 7, Issue of March 5, pp. 3979-3987, 1989
Printed in U.S.A.

Eukaryotic Signal Peptide Structure/Function Relationships

IDENTIFICATION OF CONFORMATIONAL FEATURES WHICH INFLUENCE THE SITE AND EFFICIENCY OF CO-TRANSLATIONAL PROTEOLYTIC PROCESSING BY SITE-DIRECTED MUTAGENESIS OF HUMAN PRE(ΔPRO)APOLIPOPROTEIN A-II*

(Received for publication, October 6, 1988)

Steven F. Nothwehr‡§ and Jeffrey I. Gordon‡¶‖

From the Departments of ‡Biological Chemistry and ¶Medicine, Washington University School of Medicine, St. Louis, Missouri 63110

The structural features which influence the efficiency and site of cleavage by eukaryotic signal peptidase have not been fully defined. Human pre(Δpro) apolipoprotein A-II is a useful model system for such an analysis. We have recently shown that a panel of mutants which have amino acids of varying physical-chemical properties substituted for the −1 residue (Ala^{20}) of its signal peptide exhibit cleavage at one of three potential sites or, in some cases, bidirected cleavage (Folz, R. J., Nothwehr, S. F., and Gordon, J. I. (1988) *J. Biol. Chem.* 263, 2070–2078). In this present study, a subset of the pre(Δpro)Xaa^{20}apoA-II mutants were used to identify conformational features, located COOH-terminal to the hydrophobic core domain, which regulate its cotranslational translocation and proteolytic processing. Proline residues were substituted at positions 13, 14, or 15 in ten of the original position 20 mutant preproteins in order to induce formation of a β-turn structure at various positions upstream of the cleavage site. The effects of these mutations were assessed using an *in vitro* transcription/translation/microsome processing assay. Substitution of proline at position 13 resulted in a dramatic decrease in processing efficiency in all Xaa^{20} sequence contexts, while processing progressively increased when the proline was moved "downstream" to positions 14 and 15. Elongating the hydrophobic core of the Pro^{13} mutants by 1 residue reversed this effect. NH_2-terminal sequence analysis of 34 co-translationally processed mutants revealed that the optimum distance between the site of proline introduction and the site of co-translational cleavage was 4–5 residues. However, this distance could be altered by the presence of certain amino acids at positions −1 and −3. The data suggest that several structural motifs NH_2-terminal to residues −1 and −3 influence the location, site, and efficiency of cleavage by eukaryotic signal peptidase. These include specifically the length of the hydrophobic core, the location of β-turns relative to the COOH terminus of this core domain, as well as the physical-chemical properties of amino acids at potential cleavage sites.

The co-translational translocation of most secretory proteins across the endoplasmic reticular (ER[1]) membrane requires a 15–30-residue NH_2-terminal extension or signal peptide. Several functional interactions of the signal peptide with the eukaryotic translocation apparatus have been defined. During the initial steps of synthesis of secretory proteins, the signal peptide interacts with signal recognition particle (SRP) causing a temporary elongation arrest of the ribosome-nascent chain complex at least in wheat germ cell-free translation systems (1–3). Interaction of the ribosome-nascent chain-SRP complex with the ER membrane is facilited by SRP binding to the SRP receptor (also called docking protein, Refs. 4 and 5). Subsequent co-translational translocation is thought to be initiated by release of the signal peptide from SRP followed by binding to the signal sequence receptor, an integral membrane glycoprotein (6). During translocation, the signal peptide cleavage site is recognized by signal peptidase, situated on the luminal side of the ER membrane, and co-translational proteolytic processing occurs.

While the functional roles of signal peptides have been defined, the structural characteristics which mediate their specific interactions with the translocation/processing apparatus are less well understood. Amino acid sequence comparisons of signal peptides have revealed several common features (7, 8). These include a net positive charge at their NH_2 terminus followed by a block of hydrophobic residues predicted to assume an α-helical conformation upon insertion into membrane bilayers. In eukaryotes this hydrophobic region is often predicted to end with a β-turn structure just NH_2-terminal to the cleavage site (7, 9). Helix-breaking residues such as Pro and Gly are frequently encountered at the COOH terminus of this hydrophobic core. Small, nonpolar amino acids are most likely to occupy positions −1 and −3,[2] although this constraint on the physical-chemical property of amino acids is more "stringent" at position −1 compared to −3 (10).

The relative contributions of residues located between the hydrophobic core and the −1 site to various eukaryotic signal peptide functions have not been clearly resolved. Certain conformations in this area may be required for efficient translocation across the ER membrane. For example, theoretical models of protein translocation and recent experimental evidence in *Escherichia coli* suggest that secreted polypeptides are initially inserted into the membrane as a loop structure

* This work was funded by Grant HL-18577 from the National Institutes of Health and by the Mansanto Company. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Supported by a predoctoral fellowship from the Josiah P. Macy Jr. Foundation.

‖ Established Investigator of the American Heart Association.

[1] The abbreviations used are: ER, endoplasmic reticulum; SRP, signal recognition particle; apo, apolipoprotein; DTT, dithiothreitol; SDS, sodium dodecyl sulfate.

[2] Numbering in this nomenclature is relative to the site of signal peptidase cleavage, *i.e.* residue −1 is the COOH-terminal residue of the prepeptide.

(11). Residues which participate in β-turn structures would then be required in this region. However, their role in defining the site and efficiency of cleavage by *eukaryotic* signal peptidase has not been directly tested.

We have recently examined the relative favorability of 13 different amino acids, representing a wide range of physical-chemical parameters, as potential −1 residues (12). Each of these amino acids were substituted for Ala^{20} of human pre(Δpro)apolipoprotein A-II (this 95-amino acid-long mutant apolipoprotein lacks the pentapeptide prosegment represented in the primary translation product of preproapoA-II mRNA, Ref. 13). The advantage of studying this small model preprotein is that if there is no cleavage after Xaa^{20}, alternative sites for cleavage are available (*e.g.* Gly^{18} or Ser^{15}). We found that amino acids at the −1 position are critical for defining signal peptidase cleavage and that two or more sites can compete for cleavage by the endoprotease (12). The goal of the current study was to determine the effects of increasing the β-turn propensity at various positions between the hydrophobic core and the cleavage site of human pre(Δpro)apoA-II on co-translational translocation/processing.

EXPERIMENTAL PROCEDURES[3]

RESULTS

Experimental Design and Construction of Mutants—To study the effects of β-turns in the region between the hydrophobic core and the signal peptidase cleavage site on co-translational translocation/processing, proline, a residue which often terminates regions of α-helix and β-sheet structure and frequently occurs at β-turns (27, 28), was introduced at positions 13, 14, and 15 of human pre(Δpro) apoA-II (Fig. 1). There are several attractive features of this model preprotein. First, the propensity for turn formation in this region is relatively low as predicted by the rules of Chou and Fasman (27). Second, as noted above, it has multiple potential cleavage sites (*e.g.* Gly^{18} and/or Xaa^{20}) permitting an assessment of the effects of structural alterations on processing site selection. Finally, because a panel of mutants at Xaa^{20} had been previously constructed (12), the effects of introducing residues which influence secondary structure could be evaluated in several different sequence contexts around the site of cleavage.

Fig. 3 presents predictions of what effects substitutions of Pro^{13} for Ile^{13}, Pro^{14} for Cys^{14}, and Pro^{15} for Ser^{15} have on several features of this signal peptide, *i.e.* its hydrophobicity, β-sheet probability, and β-turn probability. Chou and Fasman rules (27) predict the hydrophobic core region of the wild type apoA-II signal peptide to exhibit a high propensity for β-sheet structure and low propensity for α-helix structure (data not shown). The hydrophobic core, as defined by hydrophobicity *and* secondary structure (β-sheet probability) plots, is predicted to be progressively shortened from its COOH-terminal end as the proline substitution is moved from position 15 to 13 (*panels A* and *B* of Fig. 3). A position of increased β-turn probability (defined by Chou and Fasman in Ref. 27) also correlates with the position of the introduced proline (*panel C* of Fig. 3).[5]

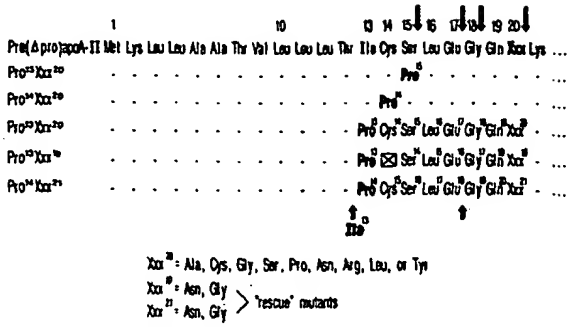

FIG. 1. **NH_2-terminal sequences of pre(Δpro)apoA-II and derivative mutants constructed for this study.** A group of 10 position 20 (Xxx^{20}) site saturation mutants (12) were selected for further mutagenesis. Proline substitutions were made at positions 13, 14, or 15 in each Xaa^{20} mutant. The $Pro^{13}Asn^{20}$ and $Pro^{13}Gly^{20}$ mutants were also used to generate deletion and insertion mutants termed $Pro^{13}Xaa^{19}$ and $Pro^{14}Xaa^{21}$, respectively. Signal peptidase cleavage sites are indicated by the *arrows*.

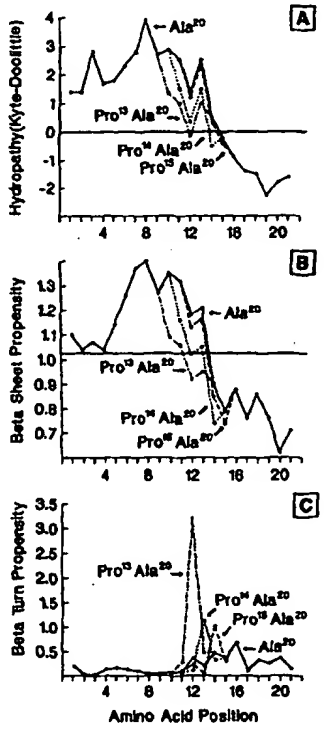

FIG. 3. **Hydrophobicity and predicted secondary structure of pre(Δpro)apoA-II and its mutant proline derivatives.** Hydrophobicity (*panel A*), β-sheet probability (*panel B*), and β-turn probability (*panel C*) are plotted for four mutants: Ala^{20} (——); $Pro^{13}Ala^{20}$ (-·-); $Pro^{14}Ala^{20}$ (····); and $Pro^{15}Ala^{20}$ (– – –). The algorithm of Kyte and Doolittle (29) was used to calculate hydrophobicity. The hydrophobicity plots were generated by averaging residue-specific hydrophobicity values for amino acids contained in a window of 4 residues. Positive and negative numbers indicate hydrophobic and hydrophilic sequences, respectively. β-sheet and β-turn probabilities were calculated according to the rules of Chou and Fasman (27) using windows of 5 and 4 residues, respectively. For example, in *panel C*, a β-turn is predicted to begin at residue 12 of the $Pro^{13}Ala^{20}$ mutant and extend for 4 residues. Values used for calculating β-sheet and β-turn probabilities are dependent on the location of a given residue within a window. As the curve for β-sheet "probability" rises above 1.05, one criteria for propagation of a β-sheet is satisfied (27).

[3] Portions of this paper (including "Experimental Procedures," Footnote 4, and Fig. 2) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are included in the microfilm edition of the Journal that is available from Waverly Press.

[5] Although the utility of applying the Chou and Fasman algorithm to peptide domains which interact with membranes such as signal sequences may be questioned (30), we feel that this represents a useful starting point for a *comparative* analysis of the predicted effects of point mutations within a given prepeptide.

The Journal of Biological Chemistry

The procedure of Mandecki (14) was used to make the Pro^{13}, Pro^{14}, and Pro^{15} substitutions in each of 10 Xaa^{20} site saturation mutants of pre(Δpro)apoA-II (Xaa^{20} = Ala, Cys, Gly, Ser, Pro, Asn, Arg, Asp, Leu, or Tyr). The construction of the $Pro^{13}Ala^{20}$ mutant is shown as an example in Fig. 2 (in the Miniprint). DNA sequence analysis of all 30 mutant cDNAs established that only the desired mutation had been made.

In Vitro Processing/Translocation Efficiency—The effects of these 30 mutations on the *efficiency* of cotranslational translocation and cleavage were analyzed by programming rabbit reticulocyte lysates containing pancreatic microsomes with mutant or wild type (preproapoA-II) mRNAs transcribed *in vitro*. Translocation was assayed by exposing the newly synthesized [^{35}S]methionine-labeled proteins to post-translationally added proteases, with or without detergent, and then immunopurifying the products. In this assay, co-translational translocation should yield a protein which is resistant to protease digestion unless its sequestration within the microsomal vesicles is abolished with detergent. The percent processing was quantitated by laser densitometric measurement of relevant bands in autoradiographs and corrected for the loss of the initiator methionine present in the signal peptide (see "Experimental Procedures" in the Miniprint).

Fig. 4 shows an experiment where Asn^{20}-derived proline mutants ($Asn^{20}Pro^{13}$, -Pro^{14}, or -Pro^{15}) plus wild type preproapoA-II were analyzed. The data reveal that the percent processing decreases progressively as the proline substitution is moved from position 15 to 13 (compare *lanes 8*, *11*, and *14*). Co-translational processing of the $Pro^{13}Asn^{20}$ is barely detectable (*lane 14*). The processing of the $Pro^{15}Asn^{20}$ (*lane 8*) mutant (42%) is significantly higher than that of the parental mutant, Asn^{20} (*lane 5* = 24%), although still lower than preproapoA-II (*lane 2*) (42% *versus* 52%). As anticipated, the translocation assay demonstrated that unprocessed preproapoA-II was *not* protected from digestion by proteases added post-translationally (*lane 3*). This was also true for 8 of the 10 "parental" Xaa^{20} and 24 of the 30 proline substitution mutants. However, this was *not* the case for the parental Asn^{20} and Gly^{20} mutants, and their corresponding Pro^{13}, Pro^{14}, and Pro^{15} "progeny." In these mutants a fraction of the *unprocessed* protein appeared to be translocated based on their protease resistance (for example, see *lanes 6*, *9*, and *12* of Fig. 4). If protease and detergent were added post-translationally to Asn^{20}, $Pro^{14}Asn^{20}$, and $Pro^{15}Asn^{20}$ mutants synthesized in the presence or absence of microsomes, complete digestion was observed (data not shown). Together these data suggest that the "insensitivity" of these unprocessed mutant proteins to protease digestion resulted from their sequestration within microsomes rather than formation of a protease-resistant structure. Additional support for the "dissociation" of co-translational translocation from signal peptidase processing in these sequence contexts comes from the fact that the unprocessed protein co-sedimented with microsomes (data not shown). Proteins which do not exhibit this "dissociation" do not co-sediment (*e.g.* wild type preproapoA-II, data not shown). For the Asn^{20}, $Pro^{14}Asn^{20}$, and $Pro^{15}Asn^{20}$ mutants, this dissociation means that the efficiency of translocation cannot be directly equated with the efficiency of processing.

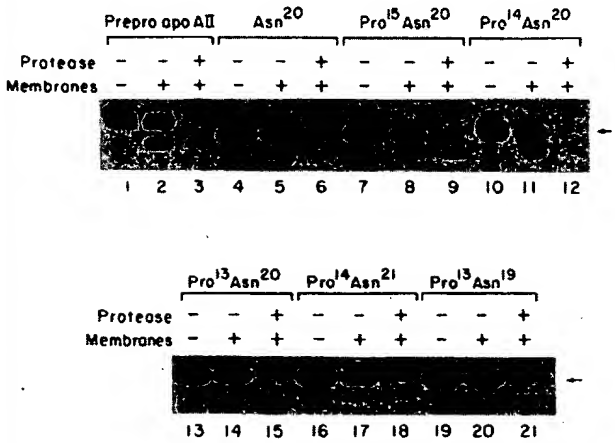

FIG. 4. ***In vitro* co-translational translocation/processing assay.** mRNAs encoding wild type preproapoA-II or pre(Δpro)apoA-II mutants were translated in reticulocyte lysates in the presence of [^{35}S]methionine. To examine the efficiency of co-translational translocation and processing, canine pancreatic microsomal membranes were added at the same time as the mRNA. Resistance to proteolysis by chymotrypsin and trypsin added post-translationally was used to establish translocation into the closed microsomal membranes. The [^{35}S]methionine-labeled polypeptides were fractionated by denaturing sodium dodecyl sulfate-polyacrylamide gel electrophoresis, and visualized by autoradiography. The positions of migration of the intact, unprocessed primary translation products are noted by *arrows*.

Experiments were performed exactly as shown in Fig. 4 for all 40 mutants. Table I lists their percent processing normalized to the parental (Xaa^{20}) mutant or to preproapoA-II (results shown in parentheses). All of the mutants exhibit significantly lower percent processing than preproapoA-II. However, comparison of the $Pro^{15}Xaa^{20}$ mutants with the Xaa^{20} mutants from which they were derived shows that signal peptidase cleavage *increases* when a Pro residue is substituted for Ser^{15} in all 10 Xaa^{20} sequence contexts. The data in Table I also demonstrate that the trend illustrated in Fig. 4 holds for all of the mutants within each position 20 sequence context, namely that movement of the proline substitution from position 15 to 13 produces a dramatic *decrease* in percent processing. Substitution of a Pro residue for Ile^{13} in each Xaa^{20} parent results in a reduction in processing to levels which are less than 15% of the Xaa^{20} parental mutant.

Inhibition of Translation by SRP—The remarkable modulation of processing efficiency by the position of the proline substitution could reflect differences in suitability of the substrate for cleavage by signal peptidase and/or differences in earlier functional interactions with the translocation appartus. To investigate the latter possibility, mutants were assayed for the relative efficiency of their interactions with SRP. SRP interaction was monitored by its ability to inhibit (*i.e.* arrest) translation of preproteins in wheat germ lysates. Table I shows the results obtained with the Asn^{20}- and Gly^{20}-proline substitution mutants. Translation arrest is expressed as the inhibition of [^{35}S]methionine-labeled mutant preprotein synthesis by exogenous SRP (the wheat germ system contains no endogenous SRP). All of the mutants show significant inhibition of translation using the conditions defined under "Experimental Procedures." In both Xaa^{20} contexts, SRP- induced arrest becomes slightly more efficient as the proline substitution moves from position 13 to 15, consistent with the observed increase in their percent processing. However, the very large differences in processing efficiency are accompanied by small differences in SRP interaction. These data raise the question of whether these mutations disturb other functional interactions, besides those involving SRP, to produce the observed reduction in processing efficiency.

Analysis of the Site of Cleavage by Signal Peptidase—To investigate whether the *position* of cotranslational processing is affected by the location of proline substitutions in various Xaa^{20} parents, the site of cleavage was identified for 22 of these mutants. This was accomplished by *in vitro* translation

TABLE I

Analysis of the site and efficiency of co-translational processing of various human pre(Δpro)apoA-II mutants

Mutation[a]	Percent cleavage after[b]				Cleavage efficiency[c]	SRP translation[d] arrest
	$Ser^{16}\downarrow$	$Glu^{17}\downarrow$	$Gly^{18}\downarrow$	$Xaa^{20}\downarrow$		
						% inhibition
Ala^{20}	0	0	0	100*		
$Pro^{15}Ala^{20}$	0	0	0	100*	118 (77)	
$Pro^{14}Ala^{20}$	0	0	82*	18	83	
$Pro^{13}Ala^{20}$	—	—	—	—*	<15	
Cys^{20}	0	0	<5*	>95		
$Pro^{15}Cys^{20}$	0	0	0*	100	109 (72)	
$Pro^{14}Cys^{20}$	0	0	86*	14	55	
$Pro^{13}Cys^{20}$	—	—	—*	—	<15	
Gly^{20}	0	0	47	53*		78
$Pro^{15}Gly^{20}$	0	0	32	68*	107 (58)	78
$Pro^{14}Gly^{20}$	0	0	100*	0	74	75
$Pro^{13}Gly^{20}$	0	<15	>85	0*	<15	62
$Pro^{14}Gly^{21}$	0	100 (Glu^{18})	0	0*	65	71
$Pro^{13}Gly^{19}$	UD	UD	UD*	UD	UD	71
Ser^{20}	0	0	74*	26		
$Pro^{15}Ser^{20}$	0	0	56*	44	127 (79)	
$Pro^{14}Ser^{20}$	0	0	100*	0	71	
$Pro^{13}Ser^{20}$	—	—	—*	—	<15	
Pro^{20}	0	0	92–95*	5–8		
$Pro^{15}Pro^{20}$	0	0	93*	7	104 (81)	
$Pro^{14}Pro^{20}$	0	0	95*	5	69	
$Pro^{13}Pro^{20}$	—	—	—*	—	<15	
Asn^{20}	0	0	100*	0		81
$Pro^{15}Asn^{20}$	0	0	100*	0	225 (72)	83
$Pro^{14}Asn^{20}$	0	0	100*	0	106	74
$Pro^{13}Asn^{20}$		<15	>85*	0	<15	73
$Pro^{14}Asn^{21}$	0	100 (Glu^{18})	0*	0	94	78
$Pro^{13}Asn^{19}$	UD	UD	UD*	UD	UD	74
Arg^{20}	0	0	100*	0		
$Pro^{15}Arg^{20}$	0	0	100*	0	108 (72)	
$Pro^{14}Arg^{20}$	0	0	100*	0	105	
$Pro^{13}Arg^{20}$	—	—	—*	—	<15	
Asp^{20}	0	0	100*	0		
$Pro^{15}Asp^{20}$	0	0	100*	0	156 (67)	
$Pro^{14}Asp^{20}$	0	0	100*	0	95	
$Pro^{13}Asp^{20}$	—	—	—*	—	<15	
Leu^{20}	<5*	0	>95	0		
$Pro^{15}Leu^{20}$	0	0	100*	0	169 (91)	
$Pro^{14}Leu^{20}$	0	0	100*	0	89	
$Pro^{13}Leu^{20}$	—	—	—*	—	<15	
Tyr^{20}	0	0	100*	0		
$Pro^{15}Tyr^{20}$	0	0	100*	0	165 (79)	
$Pro^{14}Tyr^{20}$	0	0	100*	0	96	
$Pro^{13}Tyr^{20}$	—	—	—*	—	<15	

[a] See Fig. 1 for a schematic illustration of these mutants and nomenclature.

[b] The percent cleavage after each site is based on Edman degradation of processed proteins labeled with [^{3}H] valine. Data for the site of cleavage of the parental Xaa^{20} mutants was taken from a previous report (12). A dash (—) indicates that NH_2-terminal sequence analysis was not performed. UD denotes that processing was undetectable. The asterisk (*) indicates the most likely position of co-translational cleavage as predicted by the computer program SIGSEQ2 (31). This program incorporates rules for signal peptidase cleavage based on an analysis of 161 eukaryotic signal peptides (32). SIGSEQ2 scans the NH_2-terminal residues of a preprotein and generates cleavage probability values (S values) by summing the weights contributed by residues occupying positions −13 to +2 relative to a putative cleavage site after all possible sites are considered. The signal peptidase cleavage is predicted to most likely occur at the site where the S value is the highest (31, 32). While comparisons of the predicted and observed dominant site of cleavage generally supports the power of this algorithm, a few discrepancies were noted with some of the $Pro^{14}Xaa^{20}$ mutants. First, cleavage after Gly^{18} in the Leu^{20} parent was not predicted. Second, in the Ala^{20}, Cys^{20}, Gly^{20}, and Ser^{20} sequence contexts, cleavage shifts from Gly^{18} to Xaa^{20} when the proline substitution was moved from position 14 to 15. This shift in cleavage site was not predicted to occur in the Cys^{20} sequence context. Third, in the case of the $Pro^{14}Gly^{21}$ rescue mutant, cleavage was observed after Glu^{18} while the predicted site was after Gly^{21}. These discrepancies could reflect the fact that the predictive scheme does not weight the position of the *h/c* boundary relative to potential sites of processing more heavily (see "Discussion").

[c] Calculated from laser densitometric analysis of autoradiographs taking into account the loss of one methionine (see "Experimental Procedures"). The data are expressed as the percent processing normalized to the parental Xaa^{20} mutant (set at 100%). For example, the processing efficiency of $Pro^{15}Ala^{20}$ is 118% of that for the Ala^{20} parent. A wild type preproapoA-II control was included in each experiment so that data could be compared between mutants with different Xaa^{20} residues. The percent processing of each $Pro^{15}Xaa^{20}$ mutant relative to this wild type preproapoA-II control is shown in parentheses.

[d] SRP translation arrest is expressed as the percent inhibition of translation observed in the presence of SRP as compared to that observed in the presence of an equivalent volume of SRP buffer.

of mRNAs encoding all of the $Pro^{14}Xaa^{20}$ and $Pro^{15}Xaa^{20}$ mutants, as well as the $Pro^{13}Asn^{20}$ and $Pro^{13}Gly^{20}$ constructs, in a reticulocyte lysate cell-free system containing translocation-competent canine pancreatic microsomal membranes and [^{3}H]valine. Radiolabeled apoA-II species were recovered by immunoprecipitation and the processed form further purified by preparative electrophoresis prior to automated sequential Edman degradation. The sequencing runs displayed in Fig. 5 are representative of the type of data used for analysis of co-translational processing sites. Dominant peaks of [^{3}H]valine observed at cycles 7, 11, and 18 for the $Pro^{14}Ser^{20}$ mutant (*panel A*) could have only arisen due to cleavage after Gly^{18}. In contrast, mutant $Pro^{15}Ser^{20}$ (*panel B*) shows a sequencing profile where dominant peaks of [^{3}H]valine are

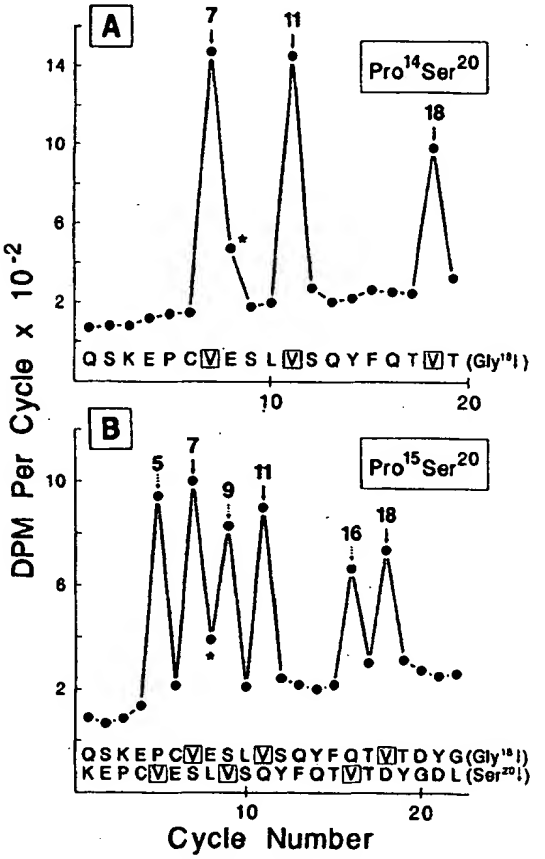

FIG. 5. **NH_2-terminal sequence analysis of co-translationally processed pre(Δpro)apoA-II mutants.** *In vitro* translation of mRNAs encoding the $Pro^{14}Ser^{20}$ and $Pro^{15}Ser^{20}$ mutants was carried out in parallel reticulocyte lysate cell-free systems containing canine pancreatic microsomal membranes and [^{3}H]valine. Radiolabeled apoA-II polypeptides were immunoprecipitated and separated by electrophoresis through 16% polyacrylamide tube gels containing 0.1% sodium dodecyl sulfate. Material passively eluted from gel slices was counted and bands corresponding to processed apoA-II protein were subjected to automated sequential Edman degradation. *Panels A* and *B* show the results obtained when 35,200 and 37,600 dpm of [^{3}H]valine-labeled $Pro^{14}Ser^{20}$ and $Pro^{15}Ser^{20}$ apoA-II were sequenced, respectively. The *solid arrows* denote peaks corresponding to cleavage after Gly^{18}, while the *dashed arrows* indicate radioactive peaks corresponding to cleavage after Ser^{20}. An *asterisk* corresponds to minor radioactive peaks due to sequencing of unprocessed apoA-II which has a valine at position 8. A [^{35}S]methionine-labeled rat apoA-IV deletion mutant, which has methionines at positions 1 and 17 (see "Experimental Procedures"), was included as an internal control in all sequencer runs to rule out the possibility of sequencing artifacts or machine malfunction. The ^{35}S data are not shown in either panel.

observed not only at cycles 7, 11, and 18 but also at cycles 5, 9, and 16. These data indicate that cleavage occurred after Ser^{20} and Gly^{18}. By measuring the amount of radioactivity delivered at cycles 16 and 18 and taking into account the repetitive yield (96%), we could calculate that 56% of the cleavage occurs after Gly^{18} and 44% after Ser^{20} in this $Pro^{15}Ser^{20}$ mutant. Note that in both sequencing runs a minor peak of [^{3}H]valine was also observed at cycle 8, corresponding to the intact, unprocessed primary translation product.

Table I lists the signal peptidase cleavage site for the 22 proline substitution mutants as well as their Xaa^{20} parents. Cleavage of the Pro^{14} and $Pro^{15}Xaa^{20}$ mutants occurred after Gly^{18} and Xaa^{20} either singly or in combination. In both of the Pro^{13} mutants sequenced, minor cleavage occurred after Glu^{17}, a site not previously documented in any of the parental pre(Δpro)apoA-II Xaa^{20} proteins (see Ref. 12 and Fig. 1). Pro^{14} and Pro^{15} mutants that had Leu, Pro, Asn, Tyr, Arg, and Asp residues at position 20 showed cleavage exclusively at Gly^{18} irrespective of the position of the increased β-turn propensity. In contrast, when prolines are substituted into Xaa^{20} parents which have dual cleavage sites Gly^{18} and Xaa^{20} (*i.e.* Ala^{20}, Gly^{20}, Cys^{20}, or Ser^{20}), the Pro^{14} mutation tends to "pull" the cleavage back to the Gly^{18} site while the Pro^{15} mutation tends to "push" the cleavage toward the Xaa^{20} site. This tendency to maintain a distance of 4–5 residues between the proline and the site of cleavage is also "obeyed" in the two Pro^{13} mutants which were evaluated, $Pro^{13}Asn^{20}$ and $Pro^{13}Gly^{20}$. As noted above, although their processing efficiency is low, both have minor sites of cleavage after Glu^{17} in addition to the "dominant" site after Gly^{18}.

Characterization of Mutants Designed to Rescue the Signal Peptide Function of Pro^{13} Containing Mutants—To further evaluate the structural feature(s) necessary for efficient signal peptidase processing, we made additional mutations in the $Pro^{13}Gly^{20}$ and $Pro^{13}Asn^{20}$ constructs. Our goal was to "rescue" the very poor processing efficiency of the Pro^{13} mutants by lengthening the hydrophobic core or shortening the distance between the putative β-turn structure and the cleavage sites. The rationale for the first alteration was that the Ile^{13} to Pro^{13} substitution was predicted to reduce the hydrophobic core length and/or the ability of the core to assume α-helical or β-sheet structures (Fig. 3), and that these changes may have compromised functional interactions with the translocation apparatus. The rationale for the second change was also straightforward: the distance between the Pro^{13} residue, and the cleavage site was reduced by 1 residue in order to test whether the distance from the putative β-turn structure to a favorable −1 residue had been too great for efficient cleavage to occur. With these considerations in mind, we added back the original Ile^{13} residue directly NH_2-terminal to the Pro^{13} residue (now at position 14) or deleted the Cys^{14} residue immediately downstream of Pro^{13}. The structures of these four mutants, named $Pro^{14}Asn^{21}$, $Pro^{14}Gly^{21}$, $Pro^{13}Asn^{19}$, $Pro^{13}Gly^{19}$, are summarized in Fig. 1.

The four mutants were subsequently analyzed for their processing/translocation efficiencies (Fig. 4 and Table I), their interactions with SRP (Table I), and their sites of co-translational cleavage (Table I). Comparison of *lanes 14, 17,* and *20* in Fig. 4 reveals that lengthening the hydrophobic core to produce the $Pro^{14}Asn^{21}$ mutant completely rescued processing efficiency; its efficiency was equivalent to that found for the $Pro^{14}Asn^{20}$ mutant (94 *versus* 106% of the Asn^{20} parent, Table I). However, deletion of Cys^{14} to reduce the distance between the predicted β turn and cleavage sites did not produce a rescued phenotype: processing of the $Pro^{13}Asn^{19}$

mutant was not detectable. Similar results were also obtained in the Gly^{20} context (Table I).

Interestingly, co-translational compartmentalization assays of $Pro^{14}Asn^{21}$ (Fig. 4, *lane 18*) revealed that most (67%) of the translocated protein was unprocessed suggesting that the extent of *rescue of translocation was even greater than that of processing efficiency*. A slight increase in SRP interaction was observed in all four of the rescue mutants over that of the $Pro^{13}Xaa^{20}$ mutants (Table I). Nonetheless, comparison of this increase in the $Pro^{14}Asn^{21}$ and $Pro^{14}Gly^{21}$ mutants to that of $Pro^{13}Asn^{19}$ and $Pro^{13}Gly^{19}$ did not reveal a consistent correlation between extent of translational arrest and rescue of processing ($Pro^{14}Asn^{21}$ and $Pro^{14}Gly^{21}$) or failure to process ($Pro^{13}Asn^{19}$ and $Pro^{13}Gly^{19}$).

Analysis of the co-translational cleavage sites in the "successful" $Pro^{14}Asn^{21}$ and $Pro^{14}Gly^{21}$ rescue mutants showed that in both cases all of the cleavage occurred after "Glu^{18}" (see Table I as well as Fig. 1). This is in contrast to the comparable Pro^{13} mutants which exhibit predominant cleavage after Gly^{19} (five amino acids downstream of the Pro) and minor cleavage after Glu^{17} (four amino acids COOH-terminal to this Pro). Thus, inclusion of the Ile^{13} residue in the two $Pro^{14}Xaa^{21}$ rescue mutants causes the distance between the Pro and the predominant cleavage site to decrease by 1 residue relative to the parental $Pro^{13}Xaa^{20}$ mutant (see Fig. 1 plus Footnote 6).

DISCUSSION

We have introduced proline substitutions to increase the β-turn propensity at various positions upstream of the cleavage site in a model preprotein, pre(Δpro)apoA-II. Our results demonstrate that the *site of cleavage* chosen by signal peptidase in these mutants is based on two factors: preservation of an optimum distance of 4–5 residues between the proline substitution and the site of co-translational processing, and the presence or absence of favorable residues at position −1. The *efficiency* of co-translational processing was dramatically reduced as the proline substitution (β-turn) was moved toward the COOH-terminal end of the hydrophobic core domain. Efficient processing could be subsequently rescued by lengthening the COOH-terminal end of the core by 1 residue suggesting the length of the core, defined in part by the location of the β-turn, is an important feature which regulates this process.

Previous studies have documented the length variations of the NH_2-terminal (*n*), hydrophobic (*h*), and COOH-terminal (*c*) regions in naturally occurring eukaryotic and prokaryotic signal peptides (7, 8, 33). The *n* and *h* regions exhibit considerable diversity in their length, while the *c* region has a much more fixed length. In eukaryotes the boundary between the 8–20-residue *h* region and the ~5 residue long *c* region is generally marked by charged or helix-breaking residues such as Pro or Gly. When eukaryotic signal peptides are aligned relative to their cleavage sites, the position of the *h/c* boundary (defined by hydrophobicity plots) typically occurs between amino acids −5 and −6 (8). Proline residues are rarely encountered between positions +1 and −3 and −7 to −10. However, proline is frequently represented at positions −4 to −6, supporting the notion that it plays a role in establishing the *h/c* boundary.

Physical analyses (39–42) and conformational energy calculations (43) of several eukaryotic and prokaryotic signal peptides suggest that in general the *h* region assumes an ordered secondary structure favoring α-helix formation in hydrophobic environments. These studies suggest that efficient signal peptide function may require the *h* region to have an uninterrupted secondary structure as well as a minimal length. Mutagenesis experiments have further aided in directly defining those structural characteristics of the hydrophobic core necessary for signal peptide function. For example, a deletion at the NH_2 terminus of the preproparathyroid hormone signal peptide that shortened its *h* region by 2 residues caused dramatic reductions in processing efficiency *in vitro* and *in vivo* (34). Deletion of residues from the hydrophobic core of the *E. coli* maltose binding protein (35) and the Lam B protein (36) blocks their export. In both cases, revertants were isolated that had extended the remaining hydrophobic core. Proline substitutions at position −7 of the yeast prepro-α-factor (37) and at position −8 of the *E. coli* Elt B (38) signal peptides severely decrease translocation *in vitro*.

While systematic experimental analysis of the essential structural characteristics of the *c* region has not been carried out (other than for residue −1), regions COOH-terminal to the cleavage site have been shown to be functionally important. Deletion of propeptides from preproapoA-II (13) and pretrypsinogen (44) results in redirection of signal peptidase cleavage to sites within the mature protein domain, suggesting that sequences COOH-terminal to the cleavage site are important in maintaining the fidelity of signal peptidase cleavage. Characterization of a propeptide deletion mutant of human preproapoA-I (24) and mature domain deletion mutants of bovine preprolactin (45) revealed that sequences beyond the signal peptidase cleavage site can affect translocation *efficiency*. Duffaud and Inouye (46) engineered point mutations in the *mature* domain of an Omp A/staphylococcal nuclease A fusion protein designed to alter predicted secondary structure in the region of the cleavage site. A reduction in predicted β-turn propensity correlated with reduced processing in *E. coli*, suggesting a role for this structure in signal peptidase I recognition/cleavage. Together these results are consistent with the hypothesis that structural compatibility between the signal peptide and the mature domain is necessary for optimal signal peptide function (45).

The above studies establish a framework for interpretation of the studies presented here which focused on the role of structural features in the *c* region and *h/c* boundary. Predictions of hydrophobicity, β-sheet propensity, and β-turn propensity (Fig. 3), suggest that introduction of proline residues at positions 13, 14, or 15 of pre(Δpro)apoA-II should result in movement of its *h/c* boundary. This movement correlates with the location of an associated increase in β-turn propensity. Thus these mutants permit an analysis of the effects of lengthening or shortening the hydrophobic core and at the same time a comparison of the relative importance of the length of the *c* region *versus* the location of favorable residues −3 and −1 in determining the site of cleavage.

Our data show that whether the cleavage site shifts COOH-terminal as the position of the *h/c* boundary shifts COOH-terminal generally depends on the Xaa^{20} context. As expected, the $Pro^{15}Xaa^{20}$ mutants which showed the most cleavage after position 20 (see Table I) had Xaa^{20} residues which were previously found (12) to have the most favorable performance as −1 amino acids (*i.e.* uncharged residues with low accessible surface areas such as Ala, Cys, Gly, and Ser). Similarly, parental Xaa^{20} mutants which show no cleavage at position 20 (*i.e.* basic, acidic, hydrophobic/aromatic amino acids such

[6] The choice of Glu^{18} over Gly^{19} as a −1 residue by signal peptidase may also partially reflect a preference for Ser^{16} over Leu^{17} to occupy position −3 in these $Pro^{14}Xaa^{14}$ species. Such a preference is supported by previous analysis of databases of naturally occurring eukaryotic signal peptides (10).

The Journal of Biological Chemistry

as Arg, Asp, Leu, and Tyr) are not forced into cleavage at position 20 by the Pro^{15} substitution.

Our analysis of how proline substitutions at positions 14 and 15 affect the site of signal peptidase cleavage in the two parental Xaa^{20} mutants (Ser^{20} and Gly^{20}) which exhibit cleavage after both Gly^{18} and Xaa^{20} suggests that the enzyme is selecting an optimal distance from the proline to the site of cleavage. The parental Ser^{20} mutant is co-translationally processed at two sites which occur 4 and 6 residues COOH-terminal to Cys^{14} and 3 and 5 residues COOH-terminal to Ser^{15} (see Fig. 1 and Table I). When proline is substituted at position 14 or 15, the distance from that proline to the processing site is either brought to a length of 4 residues (Pro^{14}) or pushed toward a length of 5 residues (Pro^{15}): *i.e.* in this Ser^{20} sequence context, the optimal number of amino acids from the proline residue to the site of cleavage appears to be 4–5 residues (see Fig. 1). Introduction of proline at position 14 in the Gly^{20} parent (which exhibits virtually equal cleavage between Gly^{18} and Gly^{20}) results in a shifting of all of the cleavage to the Gly^{18} site. However, the Pro^{15} substitution forces "only" 68% of cleavage to occur after Gly^{20}. The data suggest that signal peptidase slightly prefers cleavage 4 residues COOH-terminal of the proline rather than 5. If the proline residue is considered the first residue of the *c* region, our experimentally derived value for the optimal position of the *h/c* boundary would be between positions −5 and −6. This is in excellent agreement with the value obtained from inspection of databases of naturally occurring signal peptides (8).

Movement of the *h/c* boundary "induced" by the proline substitutions had effects not only on site selection, but also on processing efficiency. In all Xaa^{20} contexts, the Pro^{13} mutation dramatically decreased processing while a sequential increase in processing efficiency was observed as the proline substitutions were moved from the 13th through the 15th positions. For example, the site of processing of the Asn^{20} parent and its $Pro^{15}Asn^{20}$ offspring occurs after the same residue, Gly^{18}. Comparison of the processing efficiencies of both mutants demonstrates that substitution of a Pro for Ser at position −4 has the effect of more than doubling the percent processing. Since the length of the hydrophobic core (as predicted by hydrophobicity and β-sheet parameters) does not appreciably change as a result of the Pro^{15} for Ser^{15} substitution, it is likely that the differences in processing efficiency we observe between these two mutants as well as between the other Xaa^{20} parents and their $Pro^{15}Xaa^{20}$ progeny are due to the presence of a structural feature close to the cleavage site. The data are consistent with the notion that a β-turn structure relatively close to the cleavage site (as is the case with the $Pro^{15}Xaa^{20}$ mutants) presents a conformation which results in more efficient cotranslational translocation/processing than when it is further away (as is the $Pro^{13}Xaa^{20}$ mutants).

Analysis of SRP interactions with the proline substitution mutants and their Xaa^{20} parents reveals a very modest decrease in translation arrest as the *h/c* boundary moves further NH_2-terminal from the Xaa^{20} residue. Unless these differences in SRP interaction are amplified during later steps in the translocation process, it appears that the decrease in translocation/processing efficiency is largely due to a block (disruption) in a later step or steps. The potential nature of some of these disruptions is revealed by results obtained with the Asn^{20} parent and its $Pro^{14}Asn^{20}$ and $Pro^{15}Asn^{20}$ offspring. These three mutants have roughly similar translocation efficiencies but different proteolytic processing efficiencies. This unique but informative result suggests that the differences in observed processing efficiencies among these three mutants

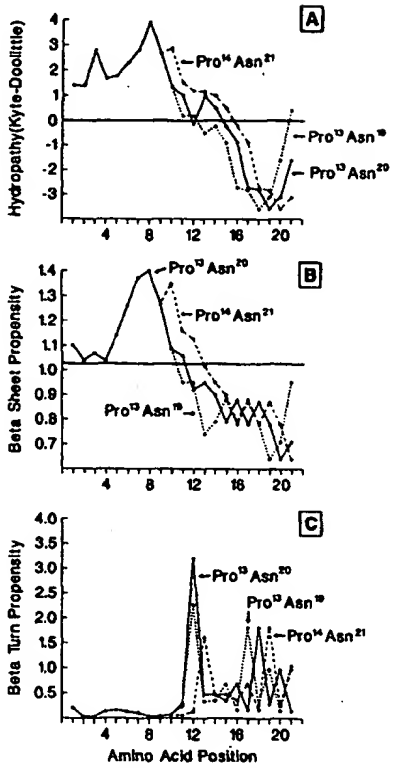

FIG. 6. **Hydrophobicity and predicted secondary structure of mutants designed to rescue the $Pro^{13}Xaa^{20}$ phenotype.** Hydrophobicity (*panel A*), β-sheet probability (*panel B*), and β-turn probability (*panel C*) are plotted for the $Pro^{13}Asn^{20}$ (——), $Pro^{14}Asn^{21}$ (– – –), and $Pro^{13}Asn^{19}$ (· · · ·) mutants as described in Fig. 3.

probably reflect their relative suitability as substrates for signal peptidase rather than differences in earlier steps.

To examine directly why the processing efficiency of the $Pro^{13}Xaa^{20}$ mutants was so low, mutants were constructed which increased the length of the hydrophobic core or decreased the distance between the *h/c* boundary and the cleavage site. These mutants were designed to distinguish whether the defect in $Pro^{13}Xaa^{20}$ mutants was due to a lack of sufficient hydrophobic core length or because the distance from the *h/c* boundary to a favorable −3, −1 site for cleavage was too great. As shown in Fig. 6, the presence of Ile^{13} in the $Pro^{14}Asn^{21}$ mutant extends the COOH-terminal end of the hydrophobic core as defined by hydrophobicity and β-sheet propensity. On the other hand, deletion of Cys^{14} in the $Pro^{13}Xaa^{19}$ mutants decreases the hydrophobicity and β-sheet probability at the COOH terminus of the *h* domain while moving the Gly residue one residue closer to the *h/c* boundary. No processing was detected for the $Pro^{13}Xaa^{19}$ mutants while the processing efficiencies of the $Pro^{14}Xaa^{21}$ mutants were rescued to levels equivalent to that of the parental (Xaa^{20} = Asn or Gly) mutants. We conclude from these data that structural alterations in the hydrophobic core of the $Pro^{13}Xaa^{20}$ mutants limited processing rather than changes in the distance from the *h/c* boundary and cleavage site. We cannot yet formally distinguish whether the Cys^{14} deletion in the $Pro^{13}Xaa^{19}$ mutants blocked our attempted rescue by perturbating the hydropathic and secondary structural features of the *h* and/or *h/c* region to the point where *translocation function* was lost, or converted it to an inactive signal peptidase substrate. We can say that steps besides interaction with SRP are probably involved since the extent of differences in processing of the

four rescue mutants do not correspond to differences in their behavior in the translational arrest assay (Table I).

In summary, these data suggest that the interruption of secondary structure and hydrophobicity of the *h* region produced by the Pro substitutions sets an optimal window of cleavage. However, this distance may be modified depending upon the physical-chemical properties of the potential −1 residue. The "reason" for the relatively fixed length of the c region is not known. One possible model is that the *h/c* boundary lines up along the interface between the hydrophobic environment of the membrane and the aqueous environment of the ER lumen. A specific length of polypeptide chain extruding from the membrane might then be optimal for recognition, interaction, and cleavage by signal peptidase. Additional rescue experiments focusing on disabled pre(Δpro)apoA-II mutants should permit further evaluation of these hypotheses.

Acknowledgments—We are indebted to our colleague Rodney Folz for his continued help and advice concerning this work, to Glen Hortin, Arnold Strauss, and Robert Heuckeroth for their insightful comments during the writing of this manuscript, and to Mark Frazier for maintaining our spinning cup protein sequencer.

REFERENCES

1. Kurzchalia, T. V., Wiedmann, M., Girshovich, A. S., Bochkareva, E. S., Bielka, H., and Rapoport, T. A. (1986) *Nature* **320**, 634–636
2. Krieg, U. C., Walter, P., and Johnson, A. E. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 8604–8608
3. Walter, P., and Blobel, G. (1981) *J. Cell Biol.* **91**, 557–561
4. Meyer, D. I., Krause, E., and Dobberstein, B. (1982) *Nature* **297**, 647–650
5. Gilmore, R., Blobel, G., and Walter, P. (1982) *J. Cell Biol.* **95**, 463–469
6. Wiedmann, M., Kurzchalia, T. V., Hartmann, E., and Rapoport, T. A. (1987) *Nature* **328**, 830–833
7. Perlman, D., and Halvorson, H. O. (1983) *J. Mol. Biol.* **167**, 391–409
8. von Heijne, G. (1985) *J. Mol. Biol.* **184**, 99–105
9. Austen, B. M. (1979) *FEBS Lett.* **103**, 308–312
10. von Heijne, G. (1983) *Eur. J. Biochem.* **133**, 17–21
11. Kuhn, A. (1987) *Science* **238**, 1413–1415
12. Folz, R. J., Nothwehr, S. F., and Gordon, J. I. (1988) *J. Biol. Chem.* **263**, 2070–2078
13. Folz, R. J., and Gordon, J. I. (1986) *J. Biol. Chem.* **261**, 14752–14759
14. Mandecki, W. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 7177–7181
15. Hanahan, D. (1985) *DNA Cloning Technology*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
16. Birnboim, H. C., and Doly, J. (1979) *Nucleic Acids Res.* **7**, 1513–1523
17. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467
18. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
19. Gordon, J. I., Budelier, K. A., Sims, H. F., Edelstein, C., Scanu, A. M., and Strauss, A. W. (1983) *J. Biol. Chem.* **258**, 14054–14059
20. Gordon, J. I., Sims, H. F., Lentz, S. R., Edelstein, C., Scanu, A. M., and Strauss, A. W. (1983) *J. Biol. Chem.* **258**, 4037–4044
21. Laemmli, U. K. (1970) *Nature* **227**, 680–685
22. Sharpe, C. R., Sidoli, A., Shelley, C. S., Lucero, M. A., Shoulders, C. C., and Baralle, F. E. (1984) *Nucleic Acids Res.* **12**, 3917–3932
23. Walter, P., and Blobel, G. (1983) *Methods Enzymol.* **96**, 682–691
24. Folz, R. J., and Gordon, J. I. (1987) *J. Biol. Chem.* **262**, 17221–17230
25. Thomas, K. A., Silverman, R. E., Jeng, I., Baglan, N. C., and Bradshaw, R. A. (1981) *J. Biol. Chem.* **256**, 9147–9155
26. Gordon, J. I., Smith, D. P., Andy, R., Alpers, D. H., Schonfeld, G., and Strauss, A. W. (1982) *J. Biol. Chem.* **257**, 971–978
27. Chou, P. Y., and Fasman, G. D. (1978) *Annu. Rev. Biochem.* **47**, 251–276
28. Smith, J. A., and Pease, L. G. (1980) *CRC Crit. Rev. Biochem.* **8**, 315–399
29. Kyte, J., and Doolittle, R. F. (1982) *J. Mol. Biol.* **157**, 105–132
30. Wallace, B. A., Cascio, M., and Mielke, D. L. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 9423–9427
31. Folz, R. J., and Gordon, J. I. (1987) *Biochem. Biophys. Res. Commun.* **146**, 870–877
32. von Heijne, G. (1986) *Nucleic Acids Res.* **14**, 4683–4690
33. von Heijne, G. (1985) *FEBS Lett.* **190**, 1–5
34. Freeman, M. W., Wiren, K. M., Rapoport, A., Lazar, M., Potts, J. T., and Kronenberg, H. M. (1987) *Mol. Endocrinol.* **1**, 628–638
35. Bankiatis, V. A., Rasmussen, B. A., and Bassford, P. J. (1984) *Cell* **37**, 243–252
36. Emr, S. D., and Silhavy, T. J. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80**, 4599–4603
37. Allison, D. S., and Young, E. T. (1988) *Mol. Cell. Biol.* **8**, 1915–1922
38. Ibrahimi, I., and Gentz, R. (1987) *J. Biol. Chem.* **262**, 10189–10194
39. Rosenblatt, M., Beaudette, N. V., and Fasman, G. D. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77**, 3983–3987
40. Shinnar, A. E., and Kaiser, E. T. (1984) *J. Am. Chem. Soc.* **106**, 5006–5007
41. Briggs, M. S., Cornell, D. G., Dluhy, R. A., and Gierasch, L. M. (1986) *Science* **233**, 206–208
42. Batenburg, A. M., Brasseur, R., Ruysschaert, J.-M., van Scharrenburg, G. J. M., Slotboom, A. J., Demel, R. A., and de Kruijff, B. (1988) *J. Biol. Chem.* **263**, 4202–4207
43. Pincus, M. R., and Klausner, R. D. (1982) *Proc. Natl. Acad. Sci. U. S. A.* **79**, 3413–3417
44. Burgess, T. L., Craik, C. S., Matsuuchi, L., and Kelly, R. B. (1987) *J. Cell Biol.* **105**, 659–668
45. Andrews, D. W., Perara, E., Lesser, C., and Lingappa, V. R. (1988) *J. Biol. Chem.* **263**, 15791–15798
46. Duffaud, G., and Inouye, M. (1988) *J. Biol. Chem.* **263**, 10224–10228
47. Inouye, S., Duffaud, G., and Inouye, M. (1986) *J. Biol. Chem.* **261**, 10970–10975

Supplementary material to

Eukaryotic Signal Peptide Structure/Function Relationships: Identification of Conformational Features which Influence the Site and Efficiency of Co-translational Proteolytic Processing by Site-Directed Mutagenesis of Human pre(ápr e)apolipoprotein AII

Steven F. Nothwehr and Jeffrey I. Gordon

EXPERIMENTAL PROCEDURES

Materials - Nuclease-treated reticulocyte lysate, T7 polymerase, RNasin, and reagents for double stranded dideoxynucleotide sequencing were purchased from Promega Biotec. Wheat germ lysate was prepared according to Anderson et al. (1983). [α-^{32}P]UTP (>3000 Ci/mmol) and L-[3,4,5-^{3}H]Valine (45 Ci/mmol) were obtained from Amersham Corp. L-[^{35}S]methionine (~1100 Ci/mmol) was purchased from Du Pont-New England Nuclear. Pancreatic microsomal membranes were prepared by the method of Walter and Blobel (1983a).

Site Directed Mutagenesis of Human Pre(ápro)apo-AII - Figure 1 summarizes the mutants that were constructed for this study. A panel of site-saturation mutants had been previously produced (12) in which 13 different amino acids were substituted for Ala^{20} of human pre(ápro)apoA-II. A subset of 10 of these "parental" mutants was chosen for substitution of a proline residue for Ile^{13}, Cys^{14} or Ser^{15} resulting in 10 additional mutants. The 41 base oligodeoxynucleotide used to direct the Pro^{13} for Ile^{13} substitution is shown in Figure 2. The 34 and 36 base oligodeoxynucleotides used to direct the Pro^{14} for Cys^{14} and the Pro^{15} for Ser^{15} substitutions were: 5'-GCTACTCCTCACCATCCCCTGCAGCCTTGAAGGACAG-3' and 5'-CTGTCCTTCAAGGGGCTGCAGATGGTGAGGAGTAGCCAC-3', respectively (the bases which mismatch with the template DNA are underlined). Two other oligodeoxynucleotides were used to construct Pro^{14}-Asn^{17} and the Pro^{15}-Asn^{18} mutants. A 39mer (5'-GTGCTACTCCTCACCATCCCCTGCAGCCTTGAAGGACAG-3') directed insertion of a Pro codon (CCC) at position 14 of the Gly^{20} and Asn^{20} mutants while a 35mer (5'-CTGTGCTACTCCTCACCCCCAGCCTTGAACGACAG-3') was used to delete the Cys^{14} codon (TGC) of Pro^{14}-Gly^{20} and Pro^{15} and Asn^{20} mutants (Figure 1).

The mutagenesis procedure of Mandecki (14) was used to substitute proline codons in each of the panel of 10 site saturation mutants. Construction of the Pro^{13}-Ala^{20} mutant is shown in Fig. 2 as an example. A pGEM2 recombinant plasmid containing a human pre(ápro)apoA-II Asn^{20} cDNA was cut with Pst I. The resulting linearized plasmid DNA (50 ng) and an oligonucleotide directing the desired mutation (20 pmol) were combined in a 50 μl "denaturation mix" that contained Tris, pH 8.0 (final concentration = 5 mM), KCl (10 mM), $MgSO_4$ (5 mM), and DTT (0.5 mM). This mixture was incubated at 100°C for 3 min and allowed to cool at room temperature for 5 minutes. The sample was then transferred to 200 μl of chilled competent E. coli strain DH5 cells (15). The transformation procedure described by Hanahan (15) was subsequently followed. The number of Ampicillin resistant colonies obtained after transformation with mixtures containing the "oligo" was typically 2-10 times the number produced using control reactions that did not contain the mutagenic oligodeoxynucleotide. The Cys^{14} + Pro^{14} and Ser^{15} + Pro^{15} mutations destroy the Pst I site so cleavage by Pst I was used as an initial screening step for identifying these substitutions in plasmid DNA preparations (16). This screening step was not possible with the Ile^{13} + Pro^{13} mutants since the Pst I site remained intact. Screening for these mutants involved double stranded dideoxy-DNA sequencing using the T7 promoter primer (17). All mutant cDNAs were subsequently completely sequenced to verify the identity and uniqueness of the Pro^{14} and Pro^{15} substitutions.

In Vitro Transcription Reactions - pGEM2 plasmids containing cDNA inserts encoding wild-type (preproapo AII) or mutant human apoA-II polypeptides were purified according to Maniatis et al. (18). Recombinant plasmids contained cDNA inserts at the Eco RI site in the appropriate orientation so that transcription of mRNA could be accomplished using the T7 promoter. The plasmids were subsequently linearized with Nhe I and in vitro run-off transcriptions performed as previously described (13). The amount of mRNA generated was calculated based on the incorporation of [^{32}P]UTP into the purified product.

In Vitro Translation and Co-translational Processing/Translocation Assays - mRNA synthesized by T7 RNA polymerase was translated in the presence of [^{35}S]methionine in a rabbit reticulocyte lysate cell-free translation system (13). To assay for the extent of co-translational processing and translocation, 30 μl reactions containing [^{35}S]methionine (37 μCi), RNAsin (40 units), canine pancreatic microsomal membranes, nuclease-treated rabbit reticulocyte lysate (10 μl), and mRNA (120 ng) were incubated for 60 min. at 25°C. To assay for co-translational translocation, the reaction products were then incubated with trypsin and α-chymotrypsin (final concentration = 250 μg/ml each) or Proteinase K (1 mg/ml) with or without Triton X-100 (0.5%) for an additional 2 hrs. at 0°C. Digestion by trypsin/α-chymotrypsin was stopped by adding aprotinin (0.15 trypsin inhibitor units) and soybean trypsin inhibitor (5 fold excess by weight) while Proteinase K digestion was stopped by addition of phenylmethylsulfonyl fluoride to 10mM. Rabbit anti-human apoA-II sera preadsorbed to protein A-Sepharose CL 4B was used to immunopurify the translation products (19). Antigen-antibody-protein A-Sepharose complexes were extensively washed (20), dissociated, and further purified by electrophoresis through 16% polyacrylamide gels containing 0.1% SDS (21). The gels were fixed, dried, and exposed to Kodak XAR-5 film at -70°C using an intensifying screen.

Autoradiograms of gels containing co-translationally cleaved, [^{35}S]methionine-labeled proteins were scanned using a LKB Ultrascan XL laser densitometer. Preproapo AII contains 2 methionine residues - one in the mature 77 amino acid plasma protein domain and 1 at the site of initiation of translation (22). To account for the loss of the initiator methionine residue after cleavage of the primary translation product, the following equation was used for quantitating processing efficiency:

$$\text{Percent Processing} = \frac{(2 \times \text{Processed Apo A-II}) \times 100}{(2 \times \text{Processed Apo A-II}) + (\text{Unprocessed Prepro A-II})}$$

Assay of Inhibition of Translation by SRP - SRP was prepared from salt-washed membranes according to Walter and Blobel (23). Thirty μl wheat germ translation mixtures contained 3 μl of SRP as eluted from DEAE-Sepharose CL-6B (or 3 μl SRP buffer, see ref. 24), globin (100 ng), preproapoA-II (75 ng), pre(ápro)AII (75 ng) or derived mutant mRNAs (75 ng). The final concentration of potassium acetate in the reactions remained at 86 mM irrespective of the presence or absence of SRP. Previous translation arrest assay studies using pre(ápro)apo AII and preproapo AII indicated that this input amount of SRP was in the linear range of a curve measuring percent inhibition of translation versus SRP concentration (24). Therefore, this "subsaturating" level was utilized to best define any differences in SRP interaction between mutants.

NH$_2$-terminal Protein Sequence Analysis - [^{3}H]Valine-labeled mutant apoA-II polypeptides, synthesized in reticulocyte lysates containing co-translationally added canine pancreatic microsomal membranes, were purified by immunoprecipitation and electrophoresis through denaturing 16% polyacrylamide/SDS tube gels (21). Tube gels were frozen, sliced into 1 mm sections, and the [^{3}H]apo AII polypeptides passively eluted into a solution of 1 mM phenylmethylsulfonyl fluoride and 1 mM benzamidine HCl. A mutant rat preapoA-IV mRNA, in which codons specifying residues 15 through 28 had been deleted[4], was translated in a parallel reticulocyte lysate mix containing [^{35}S]methionine but no microsomes. The mutant primary translation product was purified as described above except rabbit anti-rat apo-AIV sera was used (kindly supplied by Dr. M. O. Davidson, Univ. of Chicago). This labeled preapoA-IV contains [^{35}S]methionine at positions 1 and 17 and was included in every sequencing run as an internal marker for these cycles of Edman degradation. The [^{3}H] and [^{35}S] polypeptides were subjected to automated sequential Edman degradation using a 0.33 M Quadrol program and a Beckman 890C spinning cup sequenator (25,26).

[4] S. F. Nothwehr, R. J. Folz, and J. I. Gordon, manuscript in preparation.

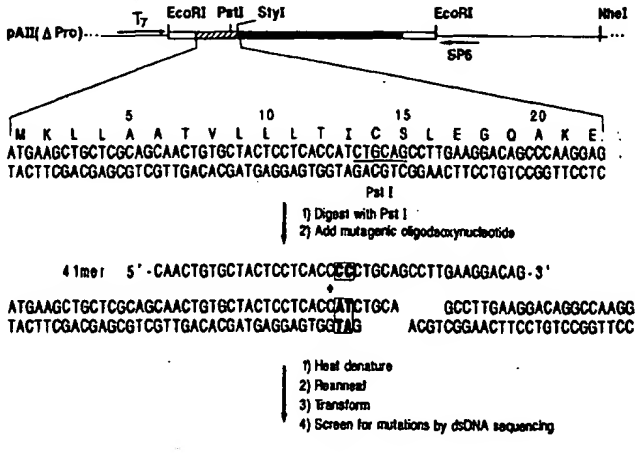

Fig. 2. Substitution of proline for isoleucine at position 13 of pre(ápro)apoA-II using oligodeoxynucleotide double-stranded break repair. Plasmid pAII(ápro)Ala^{20} (12,13) contains a human pre(ápro)apo AII cDNA. The pGEM2 plasmid sequence is represented by the thin line. Different domains of apo AII mRNA are designated by the following symbols: ▭, 5' and 3' untranslated regions; ▨, signal peptide coding domain; ▬, region encoding mature plasma protein. T_7 and SP6 refer to RNA polymerase promoter regions. Circular plasmid DNA was linearized at the unique Pst I site and combined with a 41 base mutagenic oligodeoxynucleotide in denaturation buffer (see Experimental Procedures). The mixture was incubated at 100°C, allowed to reanneal at 21°C, and then used to transform E. coli strain DH5. Colonies harboring plasmids with the desired mutation were screened by double-stranded DNA sequencing.

Proc. Natl. Acad. Sci. USA
Vol. 94, pp. 12457–12461, November 1997
Genetics

Double muscling in cattle due to mutations in the myostatin gene

ALEXANDRA C. MCPHERRON AND SE-JIN LEE*

Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, MD 21205

Communicated by Daniel Nathans, Johns Hopkins University School of Medicine, Baltimore, MD, August 26, 1997 (received for review August 12, 1997)

ABSTRACT **Myostatin (GDF-8) is a member of the transforming growth factor β superfamily of secreted growth and differentiation factors that is essential for proper regulation of skeletal muscle mass in mice. Here we report the myostatin sequences of nine other vertebrate species and the identification of mutations in the coding sequence of bovine myostatin in two breeds of double-muscled cattle, Belgian Blue and Piedmontese, which are known to have an increase in muscle mass relative to conventional cattle. The Belgian Blue myostatin sequence contains an 11-nucleotide deletion in the third exon which causes a frameshift that eliminates virtually all of the mature, active region of the molecule. The Piedmontese myostatin sequence contains a missense mutation in exon 3, resulting in a substitution of tyrosine for an invariant cysteine in the mature region of the protein. The similarity in phenotypes of double-muscled cattle and myostatin null mice suggests that myostatin performs the same biological function in these two species and is a potentially useful target for genetic manipulation in other farm animals.**

The transforming growth factor β superfamily encompasses a large group of secreted growth and differentiation factors that play important roles in regulating development and tissue homeostasis (1). We have recently described a member of this family, myostatin, that is expressed specifically in developing and adult skeletal muscle and functions as a negative regulator of skeletal muscle mass in mice (2). Myostatin null mice generated by gene targeting show a dramatic and widespread increase in skeletal muscle mass. Individual muscles in myostatin null mice weigh 2- to 3-fold more than those of wild-type mice, primarily due to an increased number of muscle fibers without a corresponding increase in the amount of fat. To pursue potential therapeutic and agricultural applications of increasing muscle mass by inhibition of myostatin activity, we have been characterizing myostatin in animals other than mice. Here we report that the myostatin gene is highly conserved among vertebrate species and that two breeds of cattle that are characterized by increased muscle mass (double muscling), Belgian Blue (3) and Piedmontese (4), have mutations in the myostatin coding sequence. These results demonstrate that the function of myostatin has been highly conserved among vertebrates.

METHODS

Cloning of Myostatin. Poly(A)-containing RNA was isolated from human (obtained from the International Institute for the Advancement of Medicine, Exton, PA), Holstein cow, sheep (Ruppersberger and Sons, Baltimore), pig (Bullock's Country Meats, Westminster, MD), White Leghorn chicken (Truslow Farms, Chestertown, MD), turkey (kindly provided by D. Boyer and D. Miller, Wampler Foods, Oxford, PA) and zebrafish (kindly provided by S. Fisher and M. Halpern, Carnegie Institution of Washington) skeletal muscle tissue as described (5). cDNA libraries were constructed in the λZAP II vector (Stratagene) according to the instructions provided by the manufacturer and screened without amplification. Rat and baboon skeletal muscle cDNA libraries and a bovine (Holstein) genomic library were purchased from Stratagene. Library screening and analysis of clones were carried out as described (5), except that the final washes were carried out in 25 mM sodium phosphate (pH 8.5), 0.5 M NaCl, 2 mM EDTA, and 0.5% SDS at 65°C.

Mapping. Fluorescence *in situ* hybridization was performed on human metaphase spreads (Bios, New Haven, CT) using a digoxigenin-labeled human genomic myostatin probe.

Sequencing of Bovine Genomic DNA. Blood from cattle was spun at 3,400 × *g* for 15 min, resuspended in 150 mM NaCl and 100 mM EDTA, and digested with 200 $\mu g \cdot ml^{-1}$ proteinase K and 1% SDS at 44°C. Semen (Select Sires, Rocky Mount, VA) was digested in 50 mM Tris (pH 8.0), 20 mM EDTA, 1% sarcosyl, 0.2 M 2-mercaptoethanol, and 200 $\mu g \cdot ml^{-1}$ proteinase K. DNAs were purified on a CsCl gradient. Exons were amplified by PCR from 1 μg genomic DNA using primer pairs 133ACM 5′-CGCGGATCCTTTGGCTTGGCGTTGCTCAAAAGC-3′ and 134ACM 5′-CGCGGATCCTTCTCATGAACACTAGAACAGCAG-3′ (exon 1), 135ACM 5′-CGCGGATCCGATTGATATGGAGGTGTTCGTTCG-3′ and 136ACM 5′-CGCGGATCCGGAAACTGGTAGTTATTTTTCACT-3′ (exon 2), and 137ACM 5′-CGCGGATCCGAGGTAGGAGAGTGTTTTGGGATC-3′ and 138ACM 5′-CGCGGATCCCACAGTTTCAAAATTGTTGAGGGG-3′ (exon 3) at 94°C for 1 min, 52°C for 2 min, and 72°C for 2 min for 40 cycles. PCR products were digested with *Bam*HI, subcloned into pBluescript, and sequenced.

Southern Blot Analysis of Mutant Sequences. One-fifth of exon 3 amplification products were electrophoresed on 2% agarose gels, blotted to nylon membranes, hybridized with ^{32}P-labeled 13-mers as described (6), and washed in 30 mM sodium citrate, 300 mM NaCl, and 0.1% SDS. Primers used were 146 ACM 5′-ATGAACACTCCAC-3′ (Holstein wild-type sequence, nucleotides 936–948), 145ACM 5′-TTGTGACAGAATC-3′ (Belgian Blue mutation, nucleotides 931–936 with 948–954), 673SJL 5′-GAGAATGTGAATT-3′ (Holstein wild-type sequence, nucleotides 1050–1062), and 674SJL 5′-GAGAATATGAATT-3′ (Piedmontese mutation, G1056A).

RESULTS AND DISCUSSION

To clone the myostatin gene from other species, cDNA libraries were constructed from RNA isolated from skeletal

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database [baboon (accession no. AF019619), bovine (accession no. AF019620), chicken (accession no. AF019621), ovine (accession no. AF019622), porcine (accession no. AF019623), rat (accession no. AF019624), turkey (accession no. AF019625), zebrafish (accession no. AF019626), and human (accession no. AF019627)].

A commentary on this article begins on page 12249.

*To whom reprint requests should be addressed. e-mail: sejin_lee@qmail.bs.jhu.edu.

muscle tissue and screened with a mouse myostatin probe corresponding to the conserved C-terminal region, which is mature, active portion of the molecule. An alignment of the predicted amino acid sequences of murine, rat, human, baboon, bovine, porcine, ovine, chicken, turkey, and zebrafish myostatin, deduced from nucleotide sequence analysis of full-length cDNA clones, is shown in Fig. 1. All of these sequences contain a putative signal sequence for secretion and a putative RXXR proteolytic processing site (amino acids 263–266) followed by a region containing the conserved C-terminal cysteine residues found in all transforming growth factor β family members (1). As seen from this alignment, myostatin is highly conserved across species. In fact, the sequences of murine, rat, human, porcine, chicken, and turkey myostatin are 100% identical in the C-terminal region following the putative proteolytic processing site, and baboon, bovine, and ovine myostatin contain only one to three amino acid differences in the mature protein. Zebrafish myostatin is considerably more diverged and is only 88% identical to the others in this region.

The high degree of sequence conservation of myostatin across species suggests that the function of myostatin has also been conserved. To determine whether myostatin plays a role in regulating muscle mass in animals other than mice, we investigated the possibility that mutations in the myostatin gene might account for the increased muscle mass observed in double-muscled livestock breeds. Double muscling, which has been observed in many breeds of cattle for the past 190 years, appears to be inherited as a single major autosomal locus with several modifiers of phenotypic expression, resulting in incomplete penetrance (7). In the most extensively studied double-muscled breed of cattle, Belgian Blue, the double muscling phenotype (Fig. 2) segregates as a single genetic locus designated muscular hypertrophy (*mh*) (8). The *mh* mutation, which is partially recessive, causes an average increase in muscle mass of 20–25%, a decrease in mass of most other organs (9, 10), and a decrease in intramuscular fat and connective tissue (11). The *mh* locus is tightly linked to markers on a region of bovine chromosome 2 (12) that is syntenic to a region of human chromosome 2 (2q32) (13) to which we had mapped the human myostatin gene by fluorescence *in situ* hybridization (data not shown).

The similarities in phenotype between the myostatin null mice and the Belgian Blue cattle breed and the similar map positions of the myostatin gene and the *mh* locus suggested the bovine homolog of myostatin as a candidate gene for the *mh* locus. To determine whether the bovine myostatin gene is mutated in the Belgian Blue breed, all three exons of the gene from the full-blood Belgian Blue bull shown in Fig. 2 were amplified by PCR, subcloned, and sequenced. The Belgian Blue myostatin coding sequence was identical to the Holstein sequence except for a deletion of nucleotides 937–947 in the third exon (Fig. 3). This 11-nucleotide deletion causes a frame-shift which is predicted to result in a truncated protein that terminates 14 codons downstream of the site of the mutation. The deletion is expected to be a null mutation because it occurs after only the first 7 amino acids of the C-terminal region, resulting in a loss of 102 amino acids (amino acids 274–375). This mutation is similar to the targeted mutation in myostatin null mice in which the entire region encoding the mature protein was deleted (2). By Southern blot analysis, using oligonucleotides corresponding to the wild-type

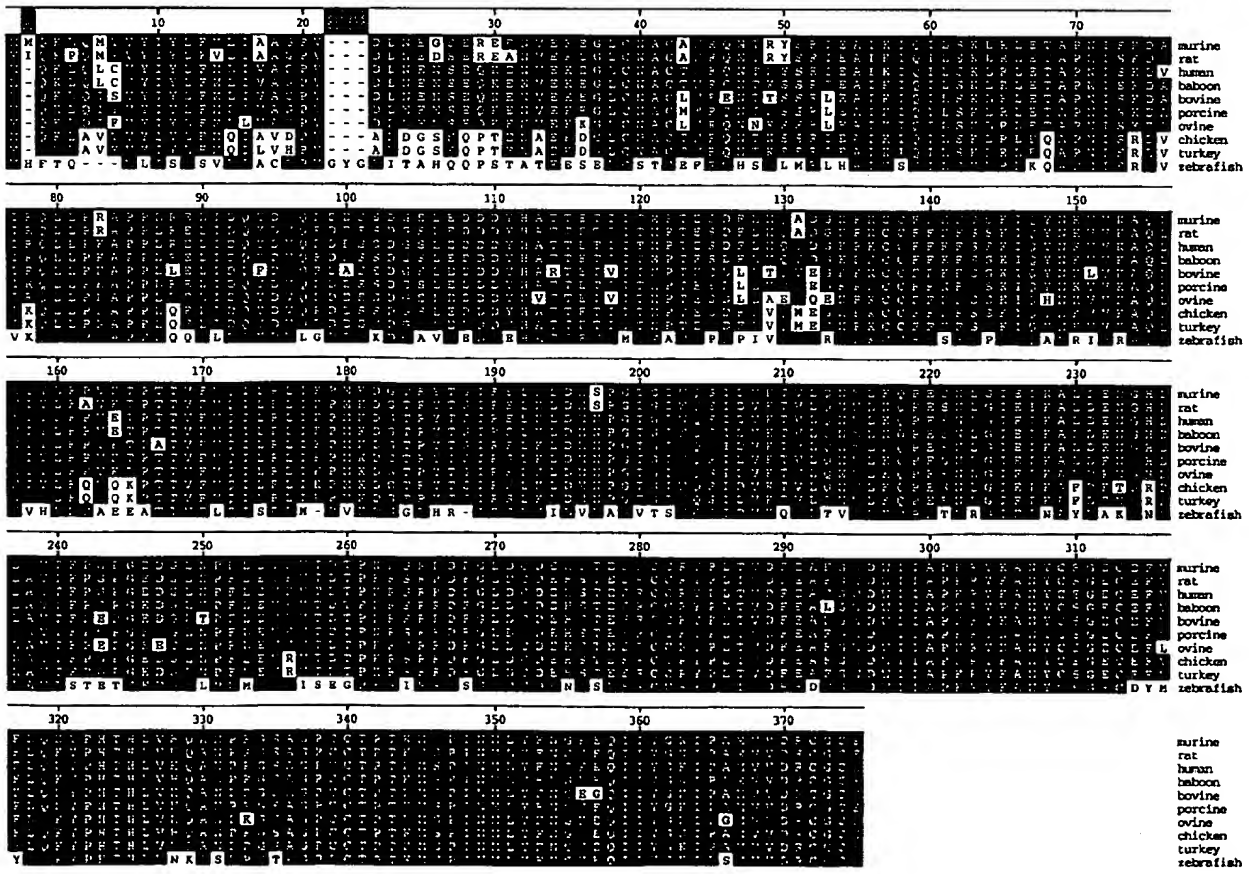

FIG. 1. Amino acid sequence alignment of murine, rat, human, baboon, bovine, porcine, ovine, chicken, turkey, and zebrafish myostatin. Shaded residues indicate amino acids matching the consensus. Amino acids are numbered relative to the human sequence. Dashed lines indicate gaps.

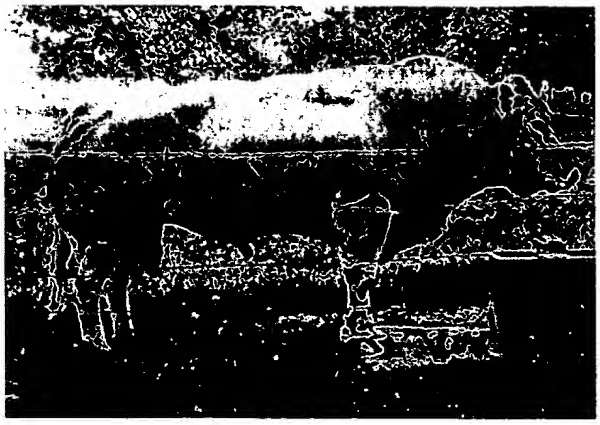

FIG. 2. A fullblood Belgian Blue bull showing the double muscling phenotype.

or mutant sequence, this mutation was found in both alleles in 14/14 fullblood Belgian Blue cattle examined (data not shown).

We also sequenced the myostatin gene in another cattle breed, Piedmontese, in which double muscling occurs at an extremely high frequency (4). The Piedmontese sequence contained 2 nucleotide changes relative to the Holstein sequence. One was a C to A transversion in exon 1, resulting in a conservative substitution of leucine for phenylalanine (amino acid 94). The second was a G to A transition in exon 3, resulting in a cysteine to tyrosine substitution in the mature region of the protein (amino acid 313) (Fig. 3). By Southern blot analysis, this mutation was found in both alleles in 10/10 double-muscled Piedmontese cattle examined. This mutation is likely to result in a complete or almost complete loss of function, as this cysteine residue is invariant not only among all myostatin sequences but also among all known members of the transforming growth factor β superfamily (1). This cysteine residue is known to be one of the amino acids involved in forming the intramolecular cystine knot structure in members of this superfamily for which the three-dimensional structure is known (14–17). Furthermore, when the corresponding cysteine in activin A (cysteine-44) was mutated to alanine, the mutant protein had only 2% of wild-type receptor binding and biological activity (18).

The similar map positions of the myostatin gene and the *mh* locus and the identification of relatively severe mutations in the myostatin gene of two different double-muscled cattle breeds suggest that these mutations are responsible for the double muscling phenotype. To further support this hypothesis, we analyzed DNA isolated from 120 individual fullblood or purebred cattle in 16 other breeds that are not classified as double-muscled (11 Angus, 11 Charolais, 10 Holstein, 10 Brown Swiss, 10 Polled Hereford, 10 Gelbvieh, 9 Simmental, 9 Jersey, 9 Guernsey, 9 Ayrshire, 7 Limousin, 4 Brahman, 4 Polled Shorthorn, 4 Red Angus, 2 Chianina, and 1 Texas

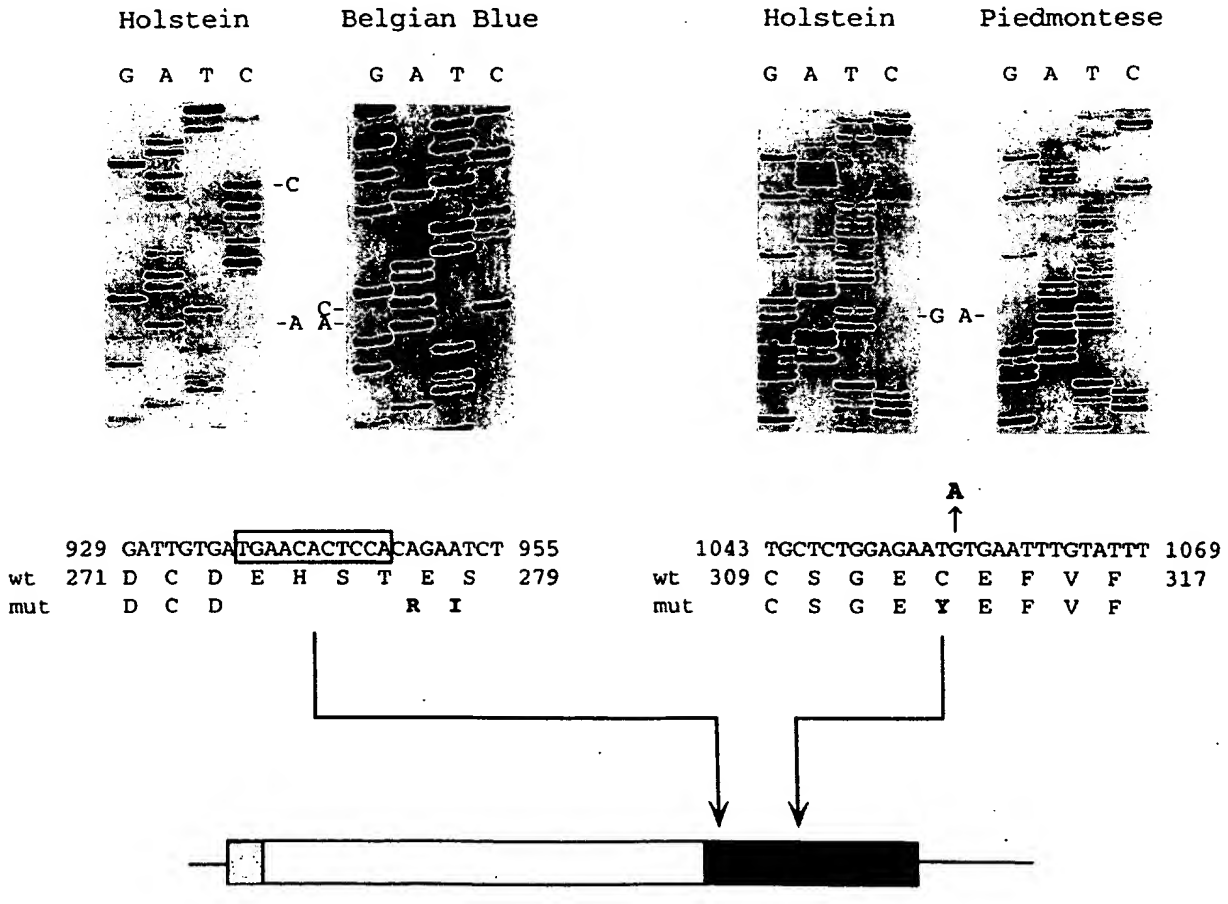

FIG. 3. Myostatin mutations in Belgian Blue (*Left*) and Piedmontese (*Right*) cattle compared with wild-type Holstein cattle. The nucleotides immediately preceding (A936) and following (C948) the Belgian Blue 11-nucleotide deletion are marked. Nucleotide and amino acid sequences are given below and numbered relative to wild type. The Belgian Blue 11-nucleotide deletion (Δ937–947) is boxed, and the Piedmontese G1056A transition is marked. Bold letters indicate nucleotide and amino acid changes. Arrows identify the locations of the mutations in the myostatin coding sequence. Shading indicates the signal sequence (gray), pro region (white) and mature C-terminal region (black).

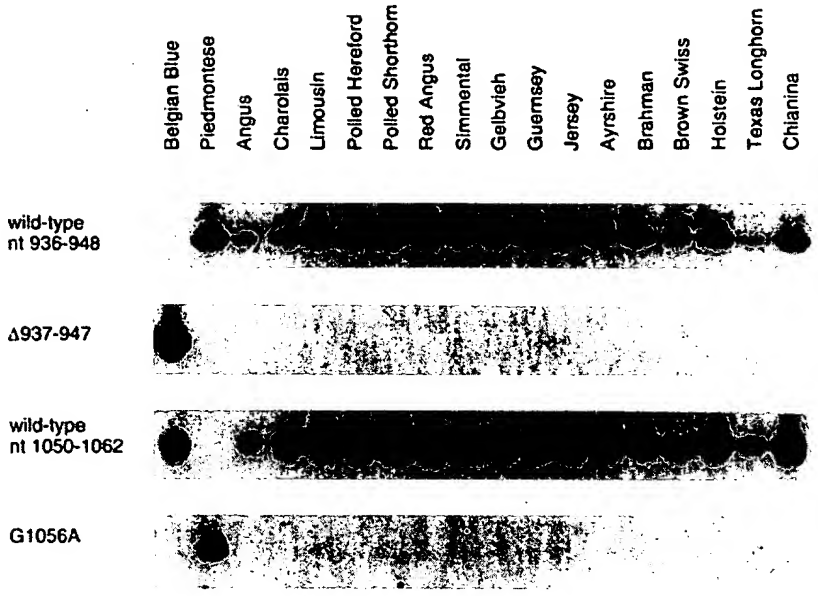

FIG. 4. Representative Southern blot hybridization showing the presence of the Belgian Blue and Piedmontese mutant sequences only in double-muscled breeds of cattle. Exon 3 PCR products were hybridized to oligonucleotide probes spanning the wild-type sequence of the region of the Belgian Blue mutation (top row), the Belgian Blue mutation Δ937–947 (second row), the wild-type sequence at nucleotide 1,056 (third row), and the Piedmontese mutant sequence at nucleotide 1,056 (bottom row). Differences in band intensity reflect differences in amounts of PCR products loaded, as judged by ethidium bromide staining (data not shown). Homozygosity for the mutations was seen only in double-muscled cattle and not in any conventional cattle as described in the text ($P < 0.001$ by χ^2).

Longhorn) for the presence of each of these mutations (Fig. 4). By Southern blot analysis, the cysteine to tyrosine substitution present in the Piedmontese breed was not detected in any of the 120 individuals. The 11-nucleotide deletion present in the Belgian Blue breed was detected in one allele of a single Red Angus non-double-muscled full-blood bull. In this regard, it has been suggested that the double muscling phenotype that is occasionally seen in many breeds may be due to a single mutation or very few mutations that migrated into many of the European breeds of cattle during the development of the modern breeds (7). Our results demonstrate that myostatin mutations which cause double muscling have occurred at least twice in cattle.

Finally, to rule out the presence of other myostatin mutations in non-double-muscled breeds, we determined the complete sequence of the myostatin coding region of 11 of these breeds (Angus, Charolais, Brown Swiss, Polled Hereford, Gelbvieh, Guernsey, Ayrshire, Limousin, Brahman, Polled Shorthorn, and Texas Longhorn). This analysis revealed only polymorphisms that were either silent changes in the coding sequences or were present in the introns and untranslated regions.

Unlike in mice, a myostatin null mutation in cattle causes a reduction in sizes of internal organs and only a modest increase in muscle mass (20–25% in the Belgian Blue breed as compared with 200–300% in myostatin-deficient mice). It is possible that cattle may be nearer to a maximal limit of muscle size after generations of selective breeding for large muscle mass, unlike mice, which have not been similarly selected. In this regard, even in cattle breeds that are not heavily muscled, the myostatin sequence contains two adjacent nonconservative amino acid differences (EG vs. KE) in the C-terminal region, compared with all other species examined. Although the functional significance of these differences is unknown, it is possible that these two changes represent a partial loss-of-function allele that became fixed in the population during many years of cattle breeding.

For agricultural applications, there are some disadvantages to double-muscled cattle, namely the reduction in female fertility, lower viability of offspring, and delay in sexual maturation (19). However, in the Belgian Blue breed, the increased muscle mass and increased feed efficiency largely offset these drawbacks (20). The fact that a null mutation in the myostatin gene in cattle results in animals that are still viable and fertile and produce high-quality meat demonstrates the potential value of producing an increase in muscle mass in other meat animals such as sheep, pig, chicken, turkey, and fish by disrupting myostatin function. Indeed, the high degree of sequence conservation in animals ranging from mammals to birds to fish suggests that the biological function of myostatin has been conserved widely throughout the animal kingdom.

We thank Dee Garrels and Chet Pennington (Lakeview Belgian Blue Ranch, Stockton, MO) for providing blood and photographs of Belgian Blue cattle. This work was supported by research grants from the Edward Mallinckrodt, Jr., Foundation and MetaMorphix, Inc. (to S.-J.L). Under an agreement between MetaMorphix, Inc. and the Johns Hopkins University, the authors are entitled to a share of sales royalty received by the University from MetaMorphix, Inc. The University, A.C.M., and S.-J.L. also own MetaMorphix stock, which is subject to certain restrictions under University policy. S.-J.L. is a consultant to MetaMorphix, Inc. The terms of this arrangement are being managed by the University in accordance with its conflict of interest policies.

1. McPherron, A. C. & Lee, S.-J. (1996) in *Growth Factors and Cytokines in Health and Disease*, eds. LeRoith, D. & Bondy, C. (JAI, Greenwich, CT), Vol. 1B, pp. 357–393.
2. McPherron, A. C., Lawler, A. M. & Lee, S.-J. (1997) *Nature (London)* **387**, 83–90.
3. Hanset, R. (1982) in *Muscle Hypertrophy of Genetic Origin and Its Use to Improve Beef Production*, eds. King, J. W. B. & Ménissier, F. (Nijhoff, The Hague, The Netherlands), pp. 437–449.
4. Masoero, G. & Poujardieu, B. (1982) in *Muscle Hypertrophy of Genetic Origin and Its Use to Improve Beef Production*, eds. King, J. W. B. & Ménissier, F. (Nijhoff, The Hague, The Netherlands), pp. 450–459.
5. Lee, S.-J. (1990) *Mol. Endocrinol.* **4**, 1034–1040.
6. Gärtner, J., Moser, H. & Valle, D. (1992) *Nat. Genet.* **1**, 16–23.

7. Ménissier, F. (1982) in *Muscle Hypertrophy of Genetic Origin and Its Use to Improve Beef Production*, eds. King, J. W. B. & Ménissier, F. (Nijhoff, The Hague, The Netherlands), pp. 387–428.
8. Hanset, R. & Michaux, C. (1985) *Genet. Sel. Evol.* **17**, 359–368.
9. Ansay, M. & Hanset, R. (1979) *Livest. Prod. Sci.* **6**, 5–13.
10. Hanset, R. (1991) in *Breeding for Disease Resistance in Farm Animals*, ed. Owen, J. B. (CAB International, Wallingford, U.K.), pp. 467–478.
11. Hanset, R., Michaux, C., Dessy-Doize, C. & Burtonboy, G. (1982) in *Muscle Hypertrophy of Genetic Origin and Its Use to Improve Beef Production*, eds. King, J. W. B. & Ménissier, F. (Nijhoff, The Hague, The Netherlands), pp. 341–349.
12. Charlier, C., Coppieters, W., Farnir, F., Grobet., L., Leroy, P. L., Michaux, C., Mni, M., Schwers, A., Vanmanshoven, P., Hanset, R. & Georges, M. (1995) *Mamm. Genome* **6**, 788–792.
13. Solinas-Toldo, S., Lengauer, C. & Fries, R. (1995) *Genomics* **27**, 489–496.
14. Daopin, S., Piez, K. A., Ogawa, Y. & Davies, D. R. (1992) *Science* **257**, 369–373.
15. Schlunegger, M. P. & Grütter, M. G. (1992) *Nature (London)* **358**, 430–434.
16. Griffith, D. L., Keck, P. C., Sampath, T. K., Rueger, D. C. & Carlson, W. D. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 878–883.
17. Mittl, P. R., Priestle, J. P., Cox, D. A., McMaster, G. Cerletti, N. & Grütter, M. G. (1996) *Protein Sci.* **5**, 1261–1271.
18. Mason, A. J. (1994) *Mol. Endocrinol.* **8**, 325–332.
19. Ménissier, F. (1982) in *Muscle Hypertrophy of Genetic Origin and Its Use to Improve Beef Production*, eds. King, J. W. B. & Ménissier, F. (Nijhoff, The Hague, The Netherlands), pp. 23–53.
20. Hanset, R., Michaux, C. & Stasse, A. (1987) *Genet. Sel. Evol.* **19**, 225–248.